ARMY RESEARCH LABORATORY

Evaluation of the ImmerVision IMV1-1/3NI Panomorph Lens on a Small Unmanned Ground Vehicle (SUGV)

by Sean Ho and Philip David

ARL-TR-6515 **July 2013**

NOTICES

Disclaimers

Army Research Laboratory

Adelphi, MD 20783-1197

ARL-TR-6515 **July 2013**

Evaluation of the ImmerVision IMV1-1/3NI Panomorph Lens on a Small Unmanned Ground Vehicle (SUGV)

Sean Ho and Philip David
Computational and Information Science Directorate, ARL

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. **REPORT DATE** *(DD-MM-YYYY)*: July 2013	2. **REPORT TYPE**: Final
3. DATES COVERED (From - To): February 2009 through March 2013	
4. TITLE AND SUBTITLE: Evaluation of the ImmerVision IMV1-1/3NI Panomorph Lens on a Small Unmanned Ground Vehicle (SUGV)	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S): Sean Ho and Philip David	**5d. PROJECT NUMBER**: 611104.H59
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES): U.S. Army Research Laboratory ATTN: RDRL-CII-A 2800 Powder Mill Road Adelphi MD 20783-1197	**8. PERFORMING ORGANIZATION REPORT NUMBER**: ARL-TR-6515
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES): U.S. Army Research, Development and Engineering Command 3071 Aberdeen Blvd. Aberdeen Proving Ground, MD 21005 Deputy Assistant Secretary of the Army 103 Army Pentagon, Washington, DC 20310-0110	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Being able to perform reconnaissance, surveillance, and target acquisition using imagers with wide fields of regard in high resolution is essential for the Army in urban warfare. In this report, we examine the performance and usability of the latest Panomorph lens, IMV1, from ImmerVision on a small unmanned ground vehicle (SUGV). The evaluation is based on the execution of two vision applications, namely, the multi-target indicator (MTI) tracker and a histogram of oriented gradients (HOG) based pedestrian detector on 60 GB or seven sets of videos collected outdoors using the IMV1 lens on a PackBot. While the IMV1 performed well with MTI with a detection range limited to 35 m, it failed to provide adequate resolution for the proper execution of the HOG-based pedestrian detector outdoors. Proper exposure compensation is a prerequisite for any outdoor task using the IMV1 due to its extremely wide field of view (FOV).

15. SUBJECT TERMS

Panomorph, omnidirectional

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified	Unclassified	Unclassified	UU	22	Sean Ho
					19b. TELEPHONE NUMBER (*Include area code*): (310) 394-3927

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

1. Introduction

Being able to perform reconnaissance, surveillance, and target acquisition using imagers with wide fields of regard in high resolution is essential for the Army in urban warfare. In order to achieve this goal, a pan-tilt-zoom (PTZ) camera has been used to allow the surveillance system to move from one area of interest (AOI) to another. A panorama can also be generated from a PTZ camera via mosaicking (*1*). A catadioptric camera that provides a 360° view in azimuth has also been used in detecting moving targets from a moving platform via the multi-sensor motion segmentation technique (*2*). Most recently, an omnidirectional camera from Point Grey Research (PGR) has been used in visual localization (*3*). This spherical camera, Ladybug, is made up of six 1/3-in charge-coupled devices (CCDs) packaged in a head unit and a base JPEG compressor with coverage of more than 75% of a full sphere. While all these cameras perform well where they are intended to be used, none is perfect. Each one has its limitation in field of view (FOV), size, or resolution when it is used on a small unmanned ground vehicle (SUGV). Consequently, we are constantly searching for a small, high-resolution lens that provides panoramic images and performs well with various vision algorithms on a SUGV platform. The latest Panomorph lens, IMV1, from ImmerVision is another example of such a lens.

The purpose of this report is to assess the performance and usability of the new IMV1-1/3NI Panomorph lens on a SUGV. The evaluation is based on the executions of two vision applications, namely, the multi-target indicator (MTI) tracker (*4*, *5*) and a histogram of oriented gradients (HOG) based pedestrian detector (*6*) on 60 GB or seven sets of videos collected using the IMV1 lens on a PackBot.

2. Background and System Setup

2.1 IMV1-1/3NI Panomorph Lens and Its Configuration

The patented ImmerVision IMV1-1/3NI panomorph lens in test has a FOV of 182° x 360°. The main advantage of this lens compared to a fisheye lens is its variable and higher pixel counts in zones of interest, especially near its periphery (*7*). Having a higher-resolution lens allows a camera to see at a longer distance. ImmerVision uses distortion as a design parameter along with anamorphic image mapping to achieve a higher pixel per degree resolution in regions closer to the peripheral (*8*, *9*). Since the operational range of most imaging systems is a function of the number of pixels on a target, having more pixels in AOIs should increase the performance accuracy of tasks such as target detection, identification, and recognition. The lens can be mounted in three different orientations: ceiling, ground, and wall (figure 1). The panomorph lens

has a standard CS mount and is optimized to be used with a camera with a 1/3-in sensor. A list of qualified cameras that have been tested to work with lens is provided by ImmerVision on its Web site (*10*). Nevertheless, all cameras on the list are Internet Protocol (IP) cameras and have large footprints, thus making them difficult to integrate to a SUGV platform. This is especially true when the lens has to be mounted low in a ground position pointing up in order to capture images in 360°. For the above reason, a 1.3-MP Chameleon color universal serial bus (USB) camera with a 1/3-in CCD from PGR was selected instead of recommended qualified cameras to host the panomorph lens. Having the advantage of a small footprint, the Chameleon camera with the IMV1 lens can be easily mounted on a PackBot (figure 2). Another advantage of using this camera is its lower power consumption, specifically, 2 W max at 5 V compared to 8 W for an average IP camera. A camera driver in a robot operating system (ROS), a development platform on the PackBot, is also available allowing the camera to capture 1280x960 images at 15 frames per second.

The disadvantage of using an USB camera to host the IMV1 is its lack of DC-Iris control, which is present on most IP cameras. This fact limited us to acquiring a version of the panomorph lens with no iris, namely, the NI version. This does not pose a problem in an indoor environment where illumination is constant. However, difficulties arise in outdoor environments where inconsistent illumination is present. In auto-exposure mode, a scene can be overexposed or underexposed due to the extra wide FOV coverage of the lens. This problem becomes prominent when the lens is positioned in a ground perspective pointing up where the sky makes up a major portion of a scene. The only way to obtain a proper exposure in this case, without a DC-Iris, is to manually adjust the shutter or/and gain on the camera body. Camera settings can be adjusted via PGR's FlyCapture graphical user interface (GUI) and software development kit (SDK). The reason that all the cameras recommended by ImmerVision are IP-based is that they provide easy integration to existing commercial surveillance and security systems.

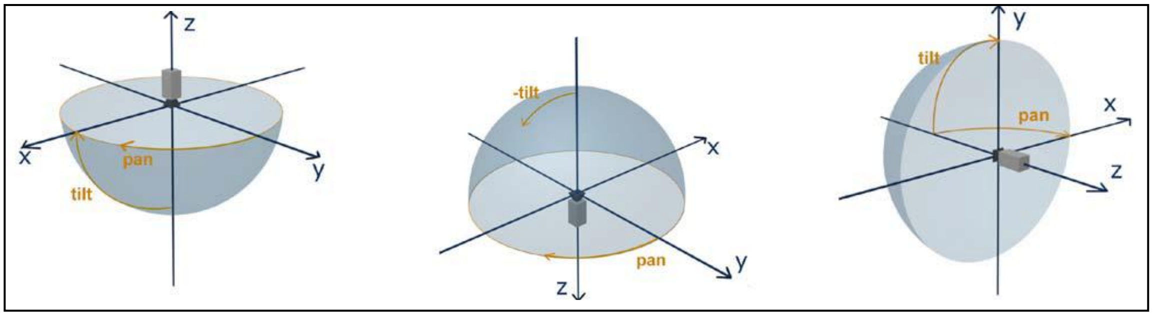


Figure 1. Three orientations for the mounting of the IMV1 lens: ceiling, ground, and wall perspective.

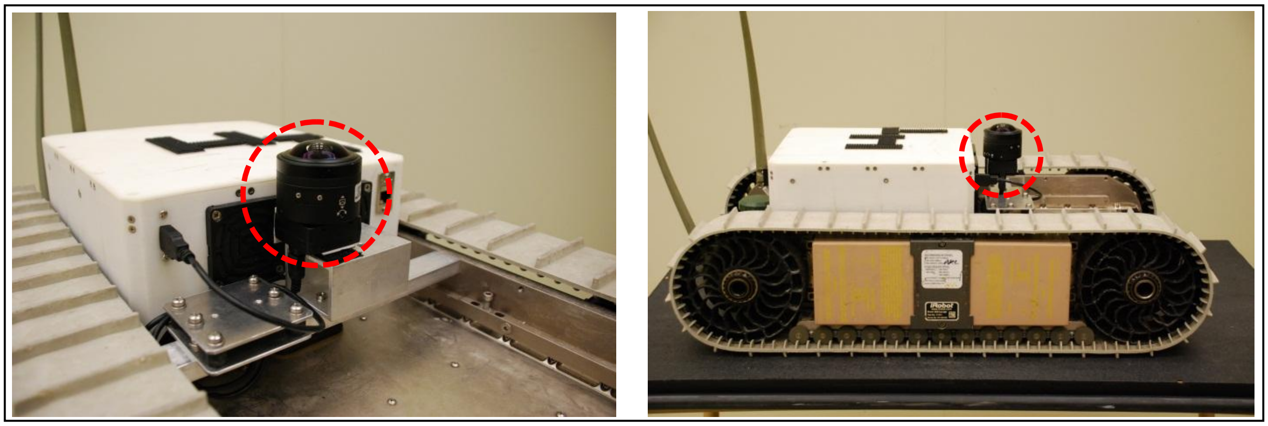


Figure 2. IMV1 with the PGR Chamaleon camera mounted on the PackBot.

2.2 PackBot Payload

The target platform for the panomorph lens is a PackBot. The payload on the PackBot runs on Ubuntu 10.10, hosts an Intel Sandy Bridge i7 quad-core central processing unit (CPU) and an Nvidia GeForce GT 545 graphics card with 144 Compute Unified Device Architecture (CUDA) cores (*11*). ROS is used on the PackBot as a software development platform for all applications. In order to obtain 360° coverage from the PackBot, the camera with the IMV1 lens is mounted in a ground perspective (figure 2). The camera is connected to the payload's USB 2.0 port. Videos are captured and recorded in ROS bag files (*12*) on a solid-state drive (SSD) via remote commands. In order not to have the payload itself be in view of the captured videos, the IMV1 is positioned only slightly above the payload in as low a position as possible to include as much of the ground plane as possible.

2.3 Software Used

2.3.1 ImmerVision Enables SDK

ImmerVision Enables is a SDK that was obtained under license from ImmerVision. The SDK allows one to call ImmerVision's shared library to visualize without distortion real-time panoramic images captured by the ImmerVision IMV1 panomorph lenses (figure 3). Input images can be unwarped and displayed in either virtual PTZ multi-views or 360° parameter views.

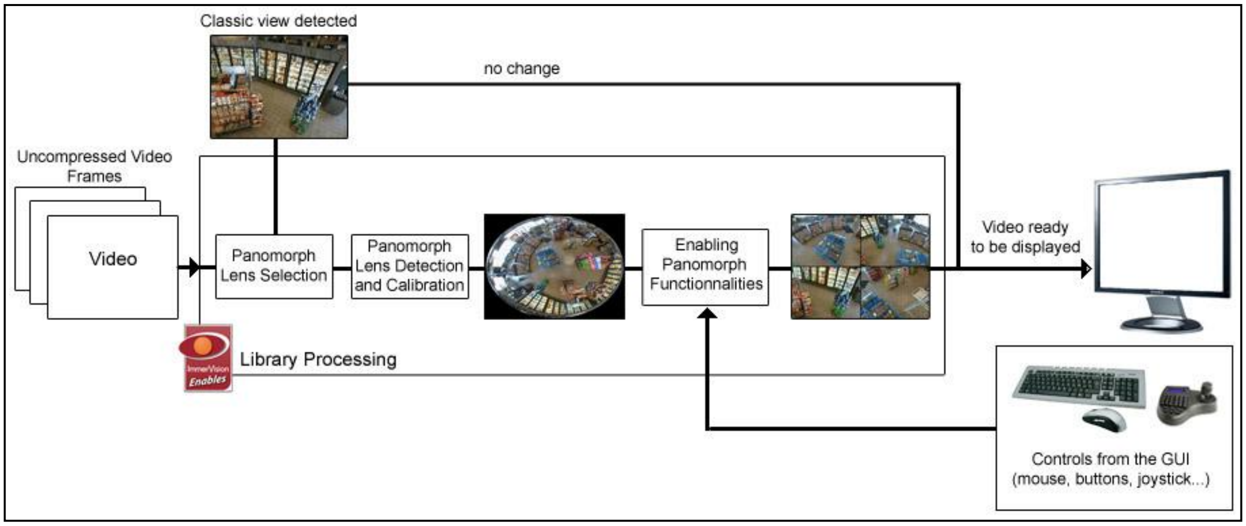
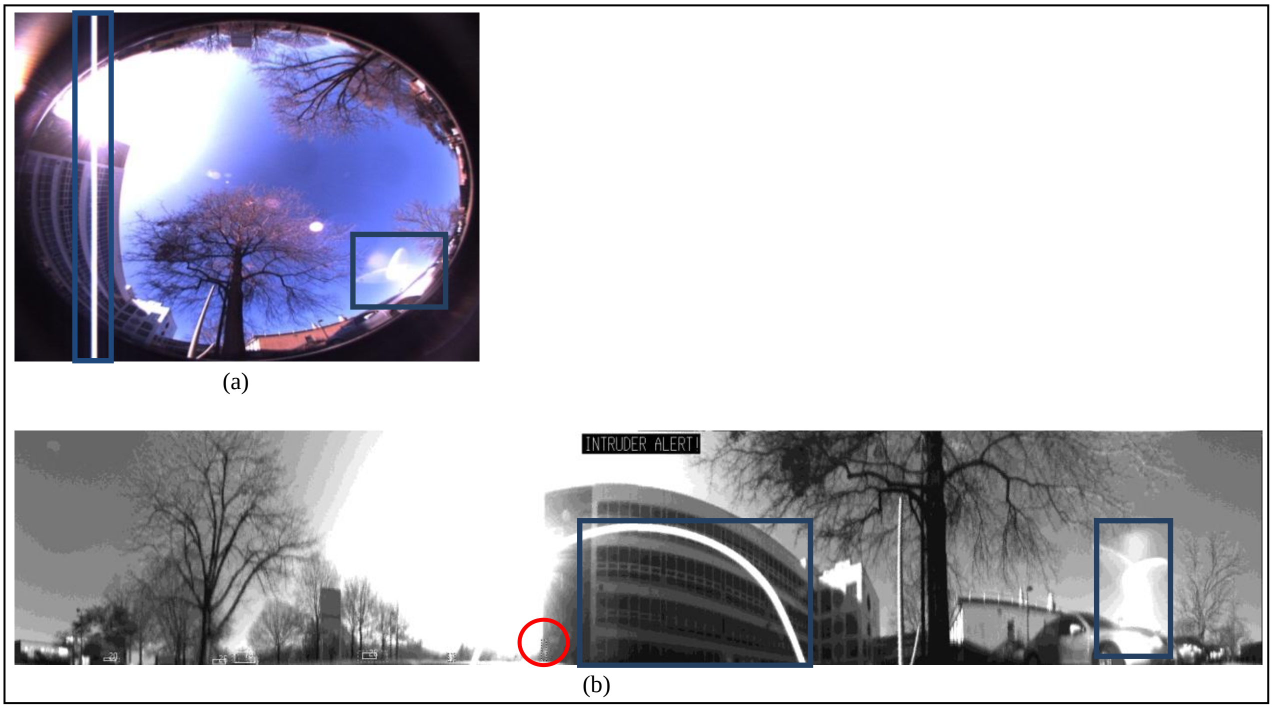


Figure 3. This flow diagram from ImmerVision shows how the SDK is used to unwarp raw images.

2.3.2 Multi-Target Indicator (MTI)

The target detection task in a surveillance system has the least requirement in input image resolution when compared to other processing tasks such as recognition and identification (*8*). An in-house developed target tracker, MTI (*4*), is used in the evaluation of the panomorph lens. Executing the MTI on captured panoramic image sequences provides a baseline assessment on the usability of the IMV1 in an outdoor environment. Targets should be detected as far away and early as possible in a detection system. The sooner targets are detected, the sooner action can be taken. The resolution of input images plays a major role in this respect. Also, the range at which targets can be detected also depends on the magnification or FOV of the sensors. A high-resolution sensor with a PTZ lens can detect and track objects kilometers away at the expense of a narrow FOV, slow focusing, and shallow depth of field. On the other hand, a panomorph lens provides an extra wide FOV at the expense of detection range. With the full 360° FOV coverage of the IMV1, targets are always in view and the only way for a target to get out of view is for it to move behind an occluding object or past its detection range so that the number of "pixels on target" falls below a threshold.

Difficulties arise in detection when low contrast exists between a target and its background. Targets moving directly at the camera may appear stationary. Furthermore, moving clutter that is common in an outdoor environment, such as smoke, tree branches, and scene illumination changes, can triggers false detections. This can be seen in video sequence 7 where steam from a heating vent on the sidewalk is detected as a moving target (figure 4b). All of these potential problems need to be addressed by the detection and tracking algorithms.



(a)

(b)

Figure 4. (a) Flares show up at two spots, shown in blue boxes, when the sun is directly in view. (b) Steam, in a red circle, is picked up as a moving object.

The detector maintains an image of the stationary components of the scene. As each new image is acquired, it is first registered to this reference image before changes between it and the reference are computed. Moving targets show up as differences between these two images. Range information is used to filter out false target detections and the remaining detections are tracked over a number of frames in order to develop high-confidence information about these targets. One advantage in running the MTI on a PackBot compared to a gas- or diesel-powered vehicle is that no image stabilization is needed to compensate for engine vibration. In order to reduce the computational load and maintain a high probability of detection, regions in a scene with high noise, clutter, or no-interest areas can be masked out.

Targets are detected as regions in the scene where the structure of the input image is changing. This is accomplished by applying a local bandpass filter to the difference between each new image and the reference image. The high-pass part of this filter reduces the effects of broad illumination changes, and the low-pass part reduces the effects of noise and other similar small changes. A threshold is then applied to the filtered difference image in order to locate regions of change before being subsampled to reduce the amount of subsequent computations without affecting the system's ability to detect targets. Blobs of potential targets are then generated.

The tracking part of the MTI finds the frame-to-frame correspondences of objects using a best-first-search approach to find the optimal correspondence of objects in the current frame with objects from past frames. The range of an object is used to limit its set of possible

correspondences to past objects based on its priori maximum speed. Possible targets are flagged when they exhibit target-like properties based on shape, size, and speed, over the last couple of frames.

2.3.3 Pedestrian Recognition

A higher requirement in image resolution or pixel count per target is expected for the successful execution of a recognition task. To evaluate the usability of the IMV1 lens on such a task, a HOG-based pedestrian detector is used (*6*). A HOG detector counts occurrences of gradient orientation as feature vectors on an image divided into a dense grid of uniformly spaced cells. The combined vectors are fed to a cascade of linear support vector machines (SVMs) for pedestrian/non-pedestrian classification. A positive result on the first classifier triggers the second and so on. Both a CPU and a graphical processing unit (GPU) version of the detector are available on the PackBot with each capable of running at 10 and 20 frames per second, respectively.

3. Data Collection

Datasets 1 through 7 were collected outdoors at the University of Maryland, College Park campus with the IMV1 mounted on the PackBot in a ground perspective. The first four sets were collected at the same location (figure 5a and b), where only pedestrians were in view, using different exposure settings. Sets 5 through 7 were taken at a different location with both pedestrians and vehicles in view (figure 5c and d). The last three sets were collected indoors with the IMV1 sitting on a work bench from a wall perspective. Videos were remotely captured as ROS bag files (*12*) and stored on a SSD on the PackBot.

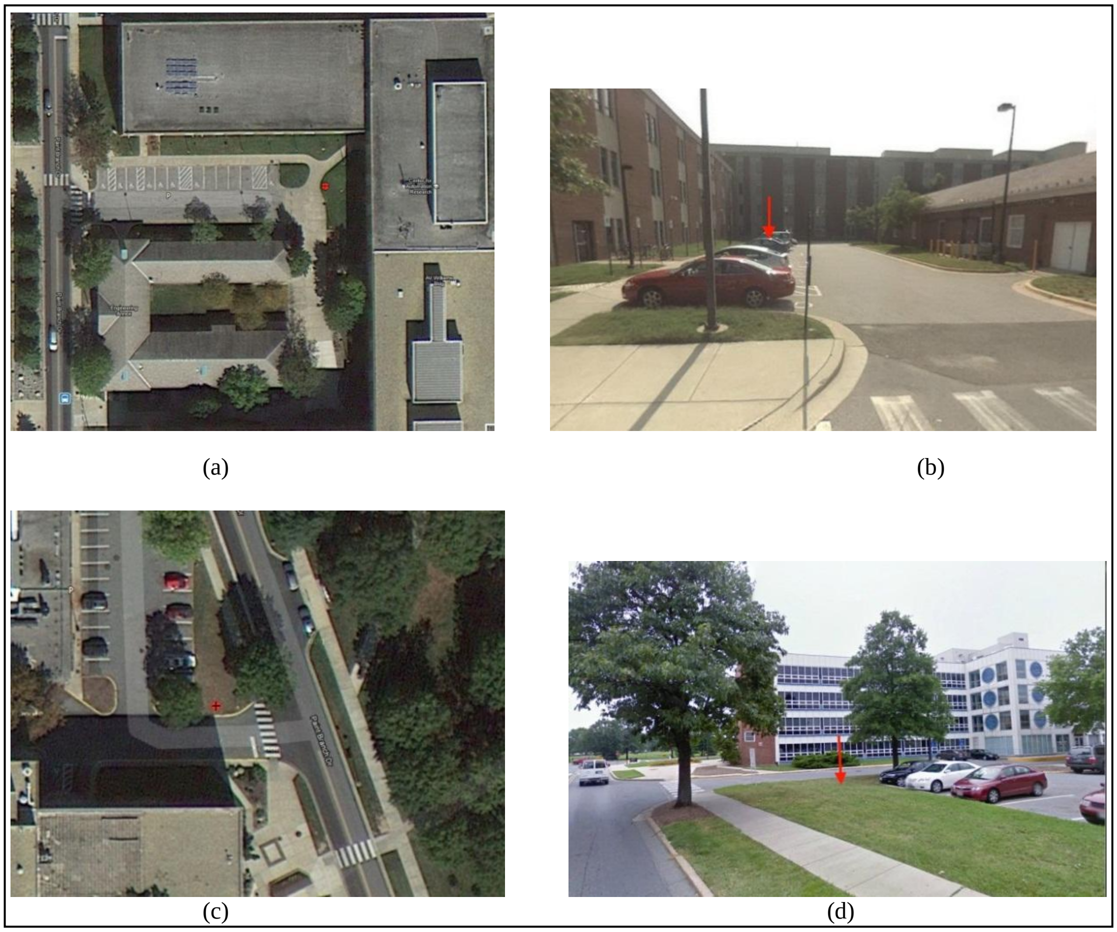


Figure 5. (a) and (b) show where the first four datesets were taken, and (c) and (d) show where the next three datasets were taken.

4. Analysis

4.1 Exposure Compensation

Images captured by the IMV1 from a ground perspective are prone to underexposure in AOIs that are usually in the periphery of an image (figure 6a). This is due to the fact that the majority of the scene in view is made up of the sky and luminance from the sky greatly skews the auto-exposure in the camera. Manual exposure adjustment or compensation is necessary in order to have the objects in the periphery of an image properly exposed (figure 6b). Without manual adjustment, objects in the peripheral are underexposed, as shown in video sequences 1 and 2. Underexposure in the AOI makes detections extremely difficult for the MTI as targets tend to blend in with the dark background. Images captured in set 3 and 4 are properly exposed after adjustments. Since DC-iris adjustment is not available on the version of IMV1 lens we have,

exposure compensation has to be made visually by adjusting the gain and shutter speed on the camera body via software.

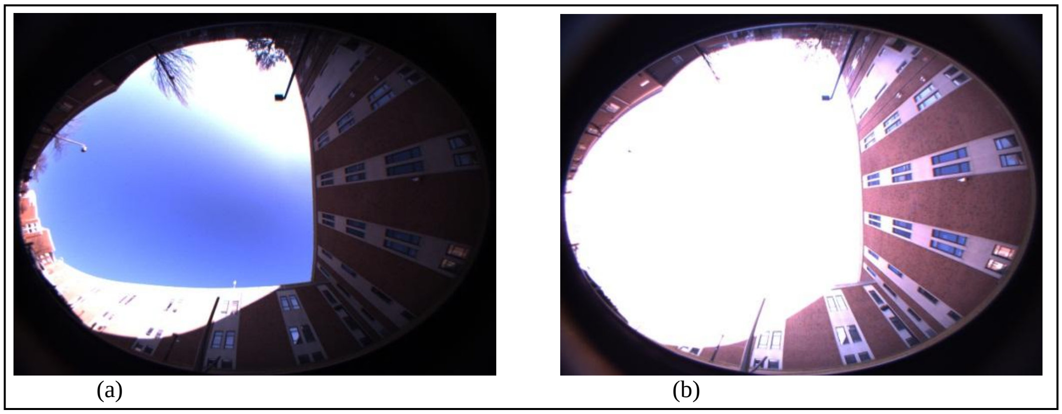
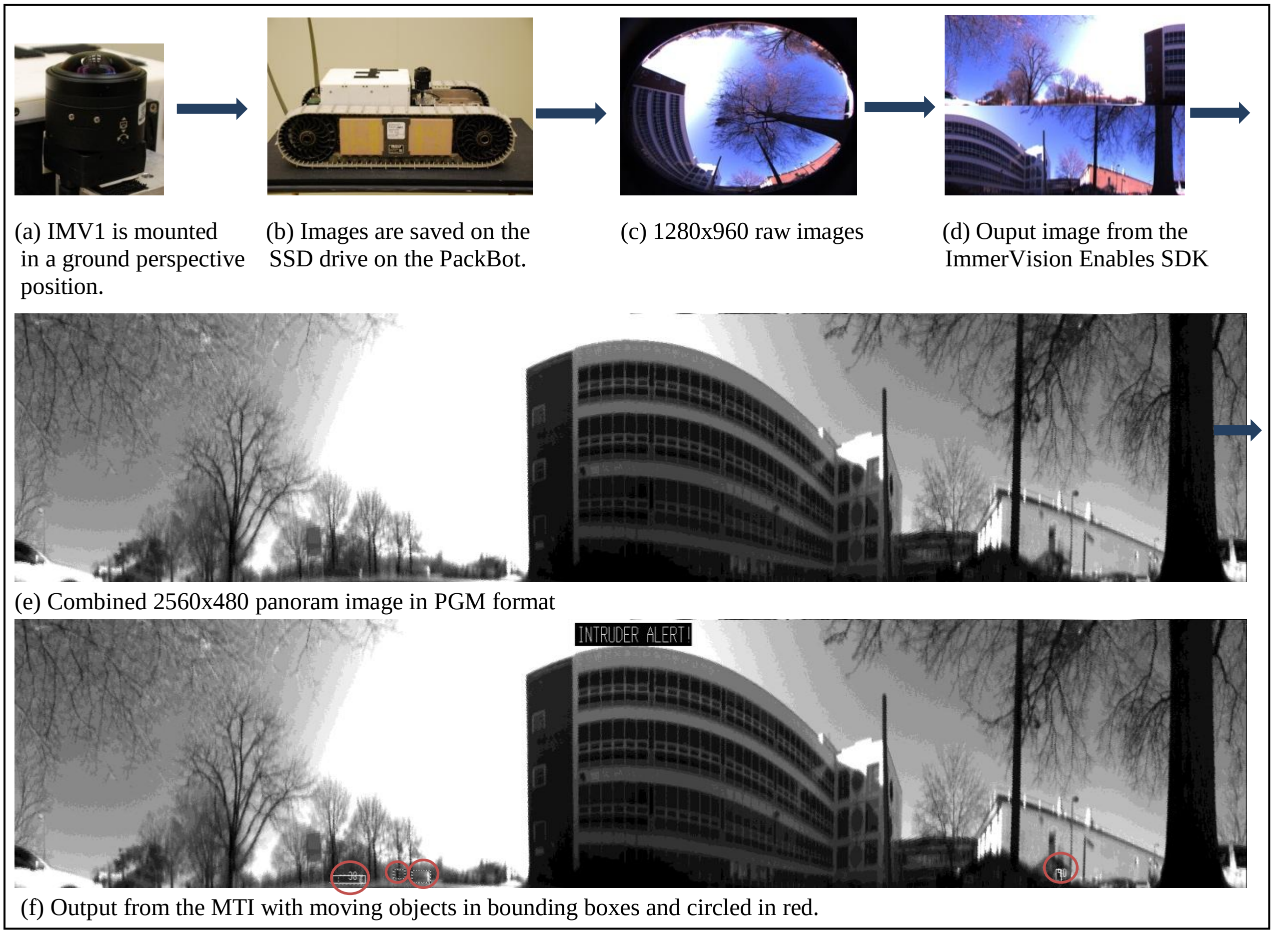


Figure 6. (a) Image is underexposed in the AOI with autoexposure. (b) Image is properly exposed in the AOI after exposure compensatoin.

4.2 Post-processing Procedures

Video sequences are saved in ROS bag files and post-processed. Each bag file is converted to individual 1280x960 jpeg images with filenames in the form of frame####.jpg, where #### depicts a frame number. These raw images are then converted to two 180° strips of 1280x480 jpeg images, namely, Frontframe####.jpg and Backframe####.jpg. This is accomplished via ImmerVision Enables SDK. Next, they are stitched and converted to a 2560x480 panorama (Combinedframe####.pgm) in portable gray map (PGM) grayscale format. These panoramas are then read in sequentially and processed by the MTI. The output from MTI is saved as mtiCombined###.pgm with bounding boxes drawn around detected moving objects. As target(s) are detected and then tracked, the system status changes from “CAUTION” to “DANGER” to “INTRUDER ALERT,” depending on the confidence level of the detection, to warn the operator. Figure 7 shows the data collection and processing steps in a flow diagram. The output images from the MTI can be played back sequentially via any image viewer that supports PGM image format.



(a) IMV1 is mounted in a ground perspective position.

(b) Images are saved on the SSD drive on the PackBot.

(c) 1280x960 raw images

(d) Ouput image from the ImmerVision Enables SDK

(e) Combined 2560x480 panoram image in PGM format

(f) Output from the MTI with moving objects in bounding boxes and circled in red.

Figure 7. This flow diagram shows how data are collected and processed.

4.3 Usability on a PackBot

For the IMV1 to be useful on a SUGV, it needs to be mounted pointing up in ground perspective in order to capture a full 360° view around it. It is a challenge to obtain proper exposure for a lens with such a wide FOV from such a position, especially in an outdoor environment. Manual shutter and gain adjustments have to be made visually on the camera to ensure an AOI is properly exposed. If it is left in auto exposure mode, the periphery of an image, where most activities occur, will be underexposed. This causes non-detections in MTI due to the blending of a person’s clothes with shadows in the background. Once proper exposure is obtained, the MTI returns surprisingly good detections with a maximum detection range of approximately 35 m for pedestrians from properly exposed video sequences 3 and 4 (figure 8). MTI also performs well with the IMV1 in tracking both vehicles and pedestrians in video sequences 5, 6, and 7 (figure 7f). In these videos, the PackBot is parked on a grassy area next to a sidewalk about 6 in higher than the road. This extra height discrepancy causes parts of a vehicle’s tires and a pedestrian’s feet out of view even though the IMV1 has a vertical FOV of 182°.

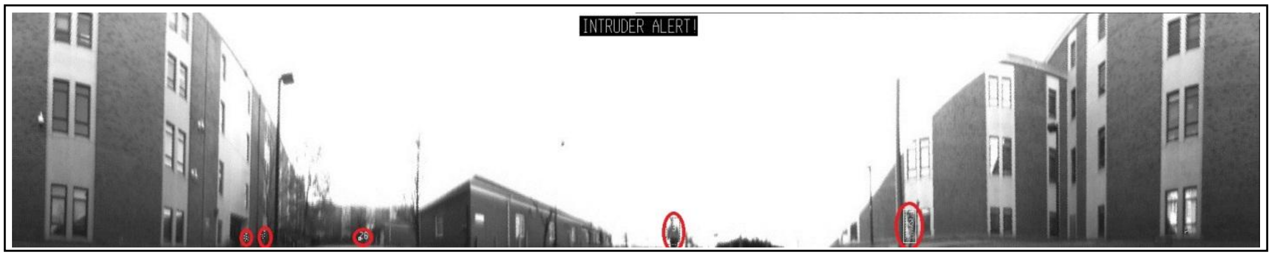


Figure 8. A sample of MTI detections of pedestrians from datasets 3 and 4. Detections are drawn in red circles.

Lens flares are present obstructing part of a scene as shown in video set 7 (figures 4a and b). A lens hood to reduce the flare is not an option due to the lens' extra wide 182° FOV; anything in front of it would be in view. This is not a problem if it is mounted from a ceiling perspective in an indoor environment. While the IMV1 works well with the MTI, the HOG pedestrian detector fails to classify any pedestrian in video sequences 1 through 7. This may be due to either the IMV1 not having enough resolution on pedestrians past a few meters away or image distortion from a ground perspective. Three short videos, 8, 9, and 10, were taken indoors with the IMV1 sitting on a work bench in a wall perspective. MTI worked well from the 180° perimeter images; however, many false positives were detected from the HOG pedestrian detector. This may be caused by the classifiers used being trained on a different camera and lens or image distortion (figure 9).

Figure 9. MTI and pedestrian detection are possible with the IMV1 in an indoor environment.

5. Conclusion

The IMV1 is the first panomorph lens incorporating a variable angular resolution in its design. The panoramic image that it produces has a higher resolution on the periphery compared to a fisheye lens. This provides better image quality for surveillance applications. IMV1 is designed to perform optimally in a ceiling mount configuration in an indoor environment where illumination is fixed and distance to objects is bounded. We evaluated the lens in an outdoor environment as its worst-case scenario. In other words, if it worked well outdoors, it should perform even better indoors. From the evaluation of the seven sets of videos taken from the IMV1 mounted in a ground perspective on a PackBot, we found it performed well in object detection and tracking with a range of approximately 35 m. Tracking results were obtained from running the MTI. The proper setting of exposure is both crucial and tricky when it is used outdoors. This is due to the extremely wide FOV of the lens from its ground position. Manual exposure compensation has to be made according to luminance in the AOI. Glares are also present on a sunny day.

While the IMV1 performed well with MTI, it did not provide adequate resolution for the proper execution of the HOG-based pedestrian detector outdoors. The resolution requirement for the execution of a recognition task is roughly 8 pixels per target versus 2 pixels per target for a detection task (*8*). While the IMV1 may not provide sufficient resolution for a recognition task, it can be used to cue other narrow FOV cameras for distant object identification. Without a 360° panorama, a narrow FOV camera would have to be continuously panned to cover the entire scene, which could significantly delay detections (*5*). The IMV1 should also work well in multisensory motion segmentation (*2*).

In conclusion, IMV1 is highly recommended to be used from a ceiling mount position in an indoor environment where luminance is evenly distributed and range is limited. It should work well indoors in real-time detection tasks and maybe in recognition tasks. It performed satisfactorily in object detection from a ground perspective in an outdoor environment with a range limited to 35 m. Proper exposure compensation is a prerequisite for any outdoor task due to its extremely wide FOV. The IMV1 failed to provide enough resolution for a recognition and identification task in an outdoor environment.

6. References

1. Ho, S.; David, P. Automatic Generation of 360 Degree Panorama from Image Sequences. *Proc. SPIE 7073, Applications of Digital Image Processing XXXI*, 70731J, 15 September 2008.

2. David, P.; Ho, S. *Multisensor Motion Segmentation*; ARL-TR-3414; U.S. Army Research Laboratory: Adelphi, MD, 2004.

3. David, P.; Ho, S. Orientation Descriptors for Localization in Urban Environments. *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pp.494–501, 25–30, September 2011.

4. David, P. Real-Time Target Acquisition for the Army's Scout Mission. *Proc. SPIE* 2739, Acquisition, Tracking, and Pointing X, 324, 7 June 1996.

5. Baran, D.; David, P. Object Tracking in an Omni-Directional Mosaic. *Proc. SPIE 7074, Advanced Signal Processing Algorithms, Architectures, and Implementations XVIII*, 707406, 3 September 2008.

6. Dalal, N.; Triggs, B, Histograms of Oriented Gradients for Human Detection. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 1*, 25–25, June 2005, 886–893.

7. ImmerVision Web site. IMV1-1/3 and IMV1-1/3NI, 360° Panomorph Lenses. http://www.immervision.com/common/security/IMV1_13_Product_Brochure.pdf (accessed 2013).

8. Thibault, S. Enhanced Surveillance System Based on Panomorph Panoramic Lenses. *Proc. SPIE* 6540, Optics and Photonics in Global Homeland Security III, 65400E, 4 May 2007.

9. Thibault S. 360 Degree Vision System: Opportunities in Transportation. *Proc. SPIE* 6758, Photonics in the Transportation Industry: Auto to Aerospace, 67580H, 26 September 2007.

10. ImmerVision Web site. Qualified Cameras & Accessories. http://www.immervision.com/en/security/security_partners/security_partners_2.php (accessed 2013).

11. Computational architecture for image processing on a small unmanned ground vehicle. *Proc. SPIE* 7798, Applications of Digital Image Processing XXXIII, 77980D, 7 September 2010.

12. ROS Web site. rosbag specifications, 26 December 2012. http://www.ros.org/wiki/rosbag (accessed 2013).

List of Symbols, Abbreviations, and Acronyms

AOI	area of interest
CCDs	charge-coupled devices
CPU	central processing unit
CUDA	Compute Unified Device Architecture
FOV	field of view
GPU	graphical processing unit
GUI	graphical user interface
HOG	histogram of oriented gradients
IP	Internet Protocol
MTI	multi-target indicator
PGM	portable gray map
PGR	Point Grey Research
PTZ	pan-tilt-zoom
ROS	robot operating system
SDK	software development kit
SSD	solid-state drive
SUGV	small unmanned ground vehicle
SVMs	support vector machines
USB	universal serial bus

NO. OF COPIES	ORGANIZATION
1 (PDF)	ADMNSTR DEFNS TECHL INFO CTR ATTN DTIC OCP
2 (PDFS)	US ARMY RSRCH LAB ATTN RDRL CIO LL TECHL LIB ATTN IMAL HRA MAIL & RECORDS MANAGEMENT
3 (PDFS)	US ARMY RESEARCH LAB ATTN RDRL-CII-A S HO P DAVID S YOUNG
1 (HC)	US ARMY RESEARCH LAB ATTN RDRL-CII B BROOME

INTENTIONALLY LEFT BLANK.